Quantitative Structure−Activity Relationships among Macrolide Antibacterial Agents: In Vitro and in Vivo Potency against *Pasteurella multocida*

James W. McFarland,*[,†,‡] Cynthia M. Berger,[§,∥] Susan A. Froshauer,[†] Shigeru F. Hayashi,[†] Scott J. Hecker,[†,⊥] Burton H. Jaynes,[†] Martin R. Jefson,[†] Barbara J. Kamicker,[†] Christopher A. Lipinski,[†] Kristin M. Lundy,[†] Catherine P. Reese,[†] and Chi B. Vu[†,#]

Central Research Division, Pfizer Inc., Groton, Connecticut 06340, and Water Technology Associates, Inc., Reading, Pennsylvania 19607

Received June 13, 1996[⊗]

Quantitative structure−activity relationships have been found among macrolide antibacterial agents in their potencies against the bacterial pathogen *Pasteurella multocida* both in vitro and in mouse infections. To obtain these relationships we measured, among other things, the pK_a's and log *P*'s of 15 known macrolides of diverse structures. Among these compounds, in vitro potency [log(1/MIC)] is a function of log *P*, log *D*, and CMR ($R = 0.86$). In vivo potency is a function of the higher pK_a, the HPLC chromatographic capacity factor log *k′*, log(1/MIC) and p*NF* ($R = 0.93$). p*NF* is defined as the negative logarithm of the fraction of neutral drug molecules present in aqueous solution at pH 7.4. The same physical properties were determined for 14 macrolides not used in developing the original QSAR models. Using the in vivo model, we calculated the mouse protection potency ranges for these new compounds. Ten estimates agreed with those observed, three were lower by a half-order of magnitude, and one was calculated to be active in the range of 15−50 mg/kg, but in fact was not active at 50 mg/kg, the highest level tested. When these new compounds were combined with the original 15, and the QSAR's updated, the new equations for the in vitro and in vivo potencies were essentially the same as those originally found. Hence, the physical properties indicated above are major determinants of macrolide antibacterial potencies.

Recently, we reported the discovery of new antibacterial repromicin derivatives, some of which are highly potent in protecting mice against lethal challenges of the Gram-negative pathogen *Pasteurella multocida*. This discovery was facilitated by the observation that common macrolides are active in vivo only when they do not exceed a specific degree of lipophilicity as estimated by HPLC methods.[1] Because macrolides are basic substances, it seemed likely that their pK_a's would also influence in vivo activity. To pursue this idea further, we measured the pK_a's, log *P*'s, and the HPLC chromatographic capacity factors (log *k*'s) of 15 known antibacterial macrolides (see Table 1).

These represent a highly diverse range of structures. There are macrolides (i) consisting of 14-membered, 15-membered, and 16-membered rings; (ii) having one, two, or three sugar moieties in several configurations; and (iii) having diverse substituents and unsaturations throughout. There is also diversity in their physical properties (see Table 2): four are dibasic, and the others monobases; their pK_a's extend over two full units, and their log *P*'s over nearly four.

We analyzed these data with multiple regression analysis and found significant new insights as to factors governing the in vitro and in vivo potencies of macrolide antibacterial agents. Later, we used these relationships to calculate the potencies of 14 macrolides that were not used in forming the model. It was found that the calculated potencies agreed well with the observed values. When the new compounds were incorporated into the data set and new models were generated, no important changes were found.

In Vitro Potency. The in vitro potencies of macrolides against *P. multocida* are expressed as minimum inhibitory concentrations (MIC's) in μg/mL and were determined as described previously.[1] For QSAR purposes, we converted these values to mmol/L and recorded them as the logarithms of their reciprocals [log(1/MIC)] in Table 2. log(1/MIC) is the dependent variable in eq 11 below, but it is also an independent variable in eq 12.

In Vivo Potency. The in vivo potencies of the macrolides are expressed as PD_{50}'s, the dose in mg/kg that protects 50% of mice against a lethal challenge of *P. multocida*.[1] Because of the large standard errors often found in determining them, and because some macrolides were not active in vivo above the highest practical dose (e.g. >150 mg/kg), we elected to use a coding system so that all 15 macrolides could be included in the analysis. The system classifies the in vivo potencies according to the dose range in which they fall. These are indicated in Table 3. Each division represents approximately a half-order of magnitude. We made one exception, azithromycin. This compound was used as the positive control and consistently gave a PD_{50} of about 16 mg/kg. This was so close to the boundary of the 2−3 classification we felt justified in assigning it the value of 2.5.

* To whom correspondence should be addressed. Tel: (860) 434-2271. Internet: james.mcfarland@snet.net.
† Central Research, Pfizer Inc.
‡ Present address: Reckon.dat Consulting, 24-2 Burr Rd, Old Lyme, CT 06371.
∥ Present address: pION Inc., 258 Harvard St, #340, Brookline, MA 02146.
§ Water Technology Associates, Inc.
⊥ Present address: Microcide Pharmaceuticals, Inc., 850 Maude Ave, Mountain View, CA 94043.
Present address: Ariad Pharmaceuticals, Inc., 26 Landsdowne St, Cambridge, MA 02139.
⊗ Abstract published in *Advance ACS Abstracts,* April 1, 1997.

S0022-2623(96)00436-0 CCC: $14.00

Table 1. In Vitro and in Vivo Activities against *Pasteurella multocida* of 15 Known Antibacterial Macrolides

compound	MIC[a] (μg/mL)	PD_{50}[b] (95% CL) (mg/kg)	MW	CAS reg. no.[c]
azithromycin	0.05	16 (13−19)	749.12	83905-01-5
erythromycylamine 11,12-carbonate	0.10	30 (23−40)	761.08	74211-76-0
erythromycylamine	0.20	6.8 (4.6−10)	735.09	26116-56-3
rosaramicin	0.20	68.0 (47−107)	581.73	35834-26-5
clarithromycin	0.39	100.0[d]	747.97	81103-11-9
repromicin	0.39	>200	565.73	56689-42-0
erythromycin	0.78	>150	734.05	114-07-8
OMT	0.78	28.0 (12−62)	597.83	61257-02-1
tilmicosin	1.56	100[e]	869.16	108050-54-0
carbomycin A	1.56	>150	842.10	4564-87-8
desmycosin	3.13	66.0 (44−99)	772.05	11032-98-7
carbomycin B	6.25	>150	826.10	21238-30-2
oleandomycin	6.25	>150	687.97	3922-90-5
desmycarosyl carbomycin A	12.50	>150	613.78	85179-34-6
tylosin	25.00	86.0 (64−119)	915.00	1401-69-0

[a] MIC: minimum inhibitory concentration. [b] PD_{50}: protective dose at which 50% of infected, treated mice survive. [c] Author provided. [d] Approximate: four to six out of 10 mice protected at 100 mg/kg. [e] Approximate: tilmicosin tested as a free base is toxic to mice at this level.

Table 2. Biological and Physical Properties of 15 Known Macrolides

compound	pK_{a1}[a]	pK_{a2}[b]	pNF[c]	pCF[c]	log P[d]	log P_i[d]	log D[e]	log k'[f]	CLOGP[g]	CMR[g]	log(1/MIC)[h]	act. class[i]
azithromycin	8.74	9.45	3.410	1.360	4.02	0.00	0.61	0.38	1.826	19.727	4.18	2.5
74211-76-0 (carbonate)	8.31	9.21	2.771	0.961	2.92	1.66	0.81	0.34	0.913	19.502	3.88	2.0
erythromycylamine	8.96	9.95	4.122	1.572	3.00	1.18	−0.32	−0.47	0.355	19.263	3.57	3.0
rosaramicin	8.79		1.407	0.017	2.19	0.00	0.85	0.63	−0.491	15.385	3.46	1.0
clarithromycin	8.99		1.601	0.011	3.16	0.00	1.57	0.63	1.276	19.394	3.28	1.0
repromicin	8.83		1.446	0.016	2.49	0.00	1.08	0.86	1.349	15.239	3.16	0.0
erythromycin	8.88		1.494	0.014	3.06	0.00	1.58	0.35	0.653	18.930	2.97	0.0
OMT	8.40		1.041	0.041	0.99	0.00	0.25	−0.17	−0.936	15.545	2.88	2.0
tilmicosin	8.18	9.56	3.007	0.847	3.80	0.00	0.79	0.26	2.502	23.093	2.75	1.0
carbomycin A	7.61		0.419	0.209	3.04	0.00	2.62	1.48	0.526	21.146	2.73	0.0
desmycosin	8.36		1.005	0.045	1.00	0.00	0.28	0.17	−0.690	19.690	2.39	1.0
carbomycin B	7.55		0.382	0.232	3.52	0.00	3.14	1.76	2.366	21.000	2.12	0.0
oleandomycin	8.84		1.455	0.015	1.69	0.00	0.43	0.01	0.596	17.672	2.04	0.0
desmycarosylcarbomycin A	8.44		1.078	0.038	0.30	0.00	0.03	−0.04	−1.688	15.263	1.69	0.0
tylosin	7.73		0.497	0.167	1.63	0.00	1.15	0.68	−0.786	23.219	1.56	1.0

[a] Lowest measured pK_a value if dibasic, or the sole pK_a value if monobasic. [b] Highest measured pK_a value if disbasic. [c] See text for details. [d] Calculated from the shift in pK_a value in the presence of rapidly stirred 1-octanol. [e] Calculated from measured pK_a, log P, and log P_i value(s). [f] Log of capacity factor (see Text for details). [g] Calculated from the MEDCHEM program (see ref 7). [h] MIC was determined as described in the Experimental Section of ref 1. [i] For a description of activity class, see text and Table 3.

Measured Descriptors. pK_a's, log P's, and log P_i's were determined by potentiometric methods as described in the Experimental Section. At 25 °C, a candidate macrolide in aqueous 0.167 M NaCl solution at low pH is titrated with 0.5 N NaOH to determine its pK_a or pK_a's. The solution is reacidified, treated with a predetermined amount of 1-octanol, and with stirring titrated again. log P and log P_i are calculated from shifts in the pK_a titration curve observed in the presence of 1-octanol.[2] Further details and the validation of the method are provided in the literature.[3,4]

One problem in analyzing this set of compounds was how to deal a mixture of mono- and dibasic compounds. Using the pK_a's directly as descriptors leaves 11 missing values in a pK_{a2} column. However, when there are two basic sites in a compound, the larger pK_a will have the greater effect on the ionization state of the molecule. Under these circumstances, it seemed reasonable to use only the *highest* pK_a as a descriptor. In the case of monbases, this will be the single pK_a, while for dibases it will be pK_{a2}.

An additional measure of lipophilicity is obtained from the capacity factors (k') of the macrolides as determined by high-performance liquid chromatography (HPLC). Retention times (t_R) of the candidate macrolides were determined as described in the Experimental Section. From this and knowledge of the column dead time (t_0), the logarithm of the capacity factor is calculated from the following formula:[5]

$$\log k' = \log(t_R/t_0 - 1) \quad (1)$$

These values are reported in Table 2. It has been shown that log k' is proportional to log P.[5]

Calculated Descriptors. The distribution coefficient, D, is the ratio of the sum of concentrations for all solute species of the macrolide of interest in the 1-octanol phase to the same for the aqueous phase. Once the pK_a's, log P's, and log P_i's are known, log D's at a selected pH can be computed. We used pH 7.4 and eqs 2 and 3 for the values recorded in Table 3. Later, we investigated some tribases; thus, we also required eq 4. These equations were derived from the basic relationship indicated above, from the various equilibrium equations involved, and by following the principles of Horváth et al.[6]

for monobases

$$\log D = \log\{P/[1 + \exp(2.3026(pK_a - 7.4))] + P_i/[1 + \exp(2.3026(7.4 - pK_a))]\} \quad (2)$$

for dibases

$$\log D = \log\{P/[1 + \exp(2.3026(pK_{a2} - 7.4))(1 + \exp(2.3026(pK_{a1} - 7.4)))] + P_i/[1 + \exp(2.3026(pK_{a1} - 7.4)) + \exp(2.3026(7.4 - pK_{a2}))] + P_{ii}/[1 + \exp(2.3026(7.4 - pK_{a1}))(1 + \exp(2.3026(7.4 - pK_{a2})))]\} \quad (3)$$

for tribases

$$\log D = \log\{P/[1 + \exp(2.3026(pK_{a3} - 7.4))(1 + \exp(2.3026(pK_{a2} - 7.4))(1 + \exp(2.3026(pK_{a1} - 7.4))))] + P_i/[1 + \exp(2.3026(7.4 - pK_{a3})) + \exp(2.3026(pK_{a2} - 7.4))(1 + \exp(2.3026(pK_{a1} - 7.4)))] + P_{ii}/[1 + \exp(2.3026(pK_{a1} - 7.4)) + \exp(2.3026(7.4 - pK_{a2}))(1 + \exp(2.3026(7.4 - pK_{a3})))] + P_{iii}/[1 + \exp(2.3026(7.4 - pK_{a1}))(1 + \exp(2.3026(7.4 - pK_{a2}))(1 + \exp(2.3026(7.4 - pK_{a3}))))]\} \quad (4)$$

In these equations P is the partition coefficient of the neutral species, P_i is that of the singly charged ion pair, P_{ii} is that of the doubly charged species, and P_{iii} is that of the triply charged species. For the compounds discussed in the present article no P_{ii}'s or P_{iii}'s were detected.

CLOGP and CMR are, respectively, the calculated log P and calculated molar refractivity obtained from the MEDCHEM program.[7] CLOGP is an alternative estimator of lipophilicity. CMR is related to a molecule's polarizability, α, which is measured in Å^3, and can be considered as an estimator of relative molecular size.[5]

Up to now, the descriptors mentioned are familiar ones that have been in the QSAR literature for many years. We now introduce two new ones. Because of their novelty we need to discuss them in greater detail.

The state of a molecule's ionization in aqueous solution plays an important role in passive transport across biological membranes. In many, if not most cases, it is the unionized molecule that is absorbed into the lipophilic layers of membranes. For lipophilic amines, such as macrolides, there is the possibility of ion pairs also penetrating biological membranes. In this case, it is highly likely that only monocharged species will be important because charge dispersion becomes increasingly difficult for dications, trications, *etc.* Hence, the lower the state of ionization, the more likely a molecule will be absorbed through the membrane.

Regardless of the number of basic sites in a molecule, all amines will have both neutral and monocharged species present in aqueous solution. From knowledge of the pK_a's, we can calculate the fraction of neutral molecules present in aqueous solution at any pH. Similarly, we can calculate the fractions of molecules bearing any number of positive charges.

To quantify this concept, we introduce two new descriptors: p*NF* and p*CF*. p*NF* is defined as the negative logarithm of the fraction of neutral molecules present in aqueous solution at pH 7.4. Similarly, p*CF* is defined as the negative logarithm of the fraction of singly charged molecules. Thus, these new descriptors are analogous to pH itself, except that p*NF* and p*CF* refer to fractions of molecular species rather than actual concentrations. The formulas used to obtain these values are given below (eqs 5−10). They are derived from the ionization equations and the expressions for the ionization constants. p*NF1* and p*CF1* are used to calculate the species for monobasic compounds; p*NF2* and p*CF2* are used for dibasic compounds; and p*NF3* and p*CF3* are used for tribasic compounds.

Table 3. Classification of Macrolides by Degree of in Vivo Potency

PD_{50} range (mg/kg)	act. class	PD_{50} range (mg/kg)	act. class
>150	0	5−15	3
50−150	1	1.5−5	4
15− 50	2		

$$pNF1 = -\log\{1/[1 + \exp(2.3026(pK_a - 7.4))]\} \quad (5)$$

$$pNF2 = -\log\{1/[1 + \exp(2.3026(pK_{a2} - 7.4)) + \exp(2.3026(pK_{a2} + pK_{a1} - 14.8))]\} \quad (6)$$

$$pNF3 = -\log\{1/[1 + \exp(2.3026(pK_{a3} - 7.4)) + \exp(2.3026(pK_{a3} + pK_{a2} - 14.8)) + \exp(2.3026(pK_{a3} + pK_{a2} + pK_{a1} - 22.2))]\} \quad (7)$$

$$pCF1 = -\log\{1 - 1/[1 + \exp(2.3026(pK_a - 7.4))]\} \quad (8)$$

$$pCF2 = -\log\{[1/[1 + \exp(2.3026(pK_{a2} - 7.4)) + \exp(2.3026(pK_{a2} + pK_{a1} - 14.8))]][\exp(2.3026(pK_{a2} - 7.4))]\} \quad (9)$$

$$pCF3 = -\log\{[1/[1 + \exp(2.3026(pK_{a3} - 7.4)) + \exp(2.3026(pK_{a3} + pK_{a2} - 14.8)) + \exp(2.3026(pK_{a3} + pK_{a2} + pK_{a1} - 22.2))]][\exp(2.3026(pK_{a3} - 7.4))]\} \quad (10)$$

These new descriptors have the somewhat confusing property (similar to the situation with pH) that the larger the value, the smaller the fraction of the species of interest.

Data Analyses. We used the FIT MULTIPLE function of RS/1 software to perform multiple linear regression analyses.[8] To compute Q^2, we used MODDE software.[9] The correlation matrix for the biological and physical properties in Table 2 is presented in Table 4.

Results. Using backward stepwise regression and the biological and physical properties in Table 2, we found the following relationship for the in vitro activity against *P. multocida*:

$$\log(1/\text{MIC}) = 0.72(\pm 0.13)\log P - 0.36(\pm 0.15)\log D - 0.15(\pm 0.05)\text{CMR} + 4.18 \quad (11)$$

$$n = 15 \quad R = 0.86 \quad s = 0.45 \quad F_{3,11} = 10.05 \quad p = 0.0017 \quad R^2 = 0.73 \quad Q^2 = 0.51$$

where n is the number of compounds employed, R is the unadjusted regression coefficient, s the standard error of the estimate, F the variance ratio with the subscript numbers indicating the degrees of freedom, and p the probability that the relationship could have occurred by chance alone. The cross-validation value, Q^2, is derived from the predictive residual sum of squares (PRESS, leave-one-out method), and is "an underestimated measure of goodness of fit...R^2 is an overestimated measure of goodness of fit".[9] A number

Table 4. Correlation Matrix of the Biological and Physical Properties of Fifteen Known Macrolides

	p$K_{a(high)}$	p*NF*	p*CF*	log *P*	log P_i	log *D*	log *k*′	CLOGP	CMR	log(1/MIC)	act. class
p$K_{a(high)}$	1.000	0.934	0.655	0.351	0.458	−0.614	−0.655	0.283	−0.122	0.667	0.589
p*NF*		1.000	0.882	0.479	0.583	−0.528	−0.552	0.359	0.088	0.693	0.733
p*CF*			1.000	0.549	0.628	−0.306	−0.304	0.384	0.345	0.584	0.771
log *P*				1.000	0.179	0.471	0.395	0.894	0.508	0.578	0.195
log P_i					1.000	−0.278	−0.308	0.051	0.069	0.466	0.574
log *D*						1.000	0.940	0.478	0.393	−0.143	−0.534
log *k*′							1.000	0.423	0.315	−0.139	−0.528
CLOGP								1.000	0.442	0.380	−0.008
CMR									1.000	−0.142	0.090
log(1/MIC)										1.000	0.628
act. class											1.000

in parentheses following a coefficient is the standard error of that coefficient.

While log *P* and log *D* both deal with the process of molecules crossing biological membranes, they are only modestly correlated with each other ($r = 0.471$, see Table 4). Nearly three-quarters of the variance in the dependent variable is accounted for by the relationship expressed in eq 11.

Again using backward stepwise regression analysis and the physical properties, but now including log(1/MIC) as an independent variable, we obtained the following relationship for in vivo potency expressed by the classification scheme described earlier:

$$\text{class} = -2.32(\pm 0.57)\text{p}K_{a(high)} + 1.26(\pm 0.31)\text{p}NF - 1.24(\pm 0.32)\log k' + 0.82(\pm 0.25)\log(1/\text{MIC}) + 17.29 \quad (12)$$

$$n = 15 \quad R = 0.93 \quad s = 0.45 \quad F_{4,10} = 14.89 \quad p = 0.000\,32 \quad R^2 = 0.86 \quad Q^2 = 0.69$$

Here we observe an even better correlation of potency; 86% of the variance in the dependent variable is accounted for by the selected descriptors. On an individual basis, p*NF* correlates highest with in vivo potency; this is followed in descending order by log(1/MIC), p$K_{a(high)}$, and finally log *k*′ (see Table 4). Taken separately or in smaller combinations, none of these descriptors correlates the data as well as when they are taken all together. One difficulty is that p*NF* is highly correlated with p$K_{a(high)}$ ($r = 0.934$). Yet the omission of either term results in an inferior model. This situation is similar to that of another frequently noticed in QSAR studies, namely the inclusion of log *P* and (log *P*)^2 terms in the same equation even when they are highly correlated with each other. On theoretical and observational grounds, this is justified by the consideration that potency does not rise continuously with increasing lipophilicity; eventually greater lipophilicity results in diminished potency.[10] The situation with p*NF* and p$K_{a(high)}$ may be somewhat akin in that these variables compensate each other for excesses at the extremes of their value ranges; after all, p*NF* is derived entirely from pK_a values. In any event, the inclusion of p*NF* and p$K_{a(high)}$ in the regression process results in a statistically significant increase in the regression sum of squares as determined by the partial *F* test.

The inclusion of the log(1/MIC) term in eq 12 is understandable because a compound's intrinsic antibacterial activity should play a role in its in vivo activity. Finally, it is gratifying to see that log *k*′ remains a significant descriptor; this was the first physical property we discovered to be associated with in vivo potency among macrolides.[1]

To confirm these results, we used eq 12 to calculate the in vivo potency classes of 14 16-membered ring macrolides for which we had the requisite data. The macrolides are the repromicin and desmycosin derivatives shown in Table 5 and are numbered the same as they were when originally reported.[1] A compound is deemed to fit the model if its PD_{50} falls between the limits of the calculated class or its 95% confidence limits overlap that range. Table 5 summarizes the data used and the results of these studies. The PD_{50}'s of 10 compounds fit the model; the potencies of **19**, **27**, and **35** were underestimated by one class (a half-order of magnitude); and **39** was calculated to be active in the 15−50 mg/kg range but was not active at 50 mg/kg, the highest dose tested.

Next, we combined the data for the 14 new macrolides with the original 15 that were used to generate eqs 11 and 12, and performed regression analyses on the new data set. Using log(1/MIC) as the dependent variable, the physical properties as the independent variables and backward stepwise regression, we found eq 13 to be the best model for in vitro potency. The descriptors and their signs for this model are the same as before, but the *R* is somewhat inferior compared to that associated with eq 11.

$$\log(1/\text{MIC}) = 0.61(\pm 0.11)\log P - 0.31(\pm 0.09)\log D - 0.14(\pm 0.04)\text{CMR} + 4.27 \quad (13)$$

$$n = 29 \quad R = 0.77 \quad s = 0.46 \quad F_{3,25} = 12.34 \quad p = 0.000\,038 \quad R^2 = 0.60 \quad Q^2 = 0.47$$

Compound **39** had no PD_{50}, thus the analysis of in vivo potency proceeded with the other 28. Incorporating log(1/MIC) with the physical properties as independent variables, we analyzed "class". Backward stepwise regression furnished eq 14 as the best model. The descriptors and their signs are the same as for eq 12; *R* was not significantly different; Q^2 was even superior.

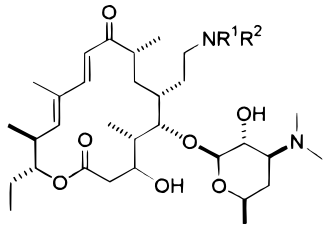
$$\text{class} = -3.01(\pm 0.55)\text{p}K_{a(high)} + 1.84(\pm 0.26)\text{p}NF - 1.24(\pm 0.30)\log k' + 0.76(\pm 0.21)\log(1/\text{MIC}) + 22.59 \quad (14)$$

$$n = 28 \quad R = 0.92 \quad s = 0.58 \quad F_{4,23} = 34.00 \quad p = 0.000\,000\,003 \quad R^2 = 0.86 \quad Q^2 = 0.78$$

These results with additional macrolides suggest that the relationships uncovered here are robust.

Discussion. Despite the fact that macrolides have been characterized as lipophilic amines,[11] there is scant data in the literature indicating just how lipophilic these compounds are, especially in direct comparison to each

Table 5. Biological and Physical Properties of 14 Repromicin and Desmycosin Derivatives



compd[a]	NR^1R^2	pK_{a1}[b]	pK_{a2}[c]	pK_{a3}[d]	pNF[e]	pCF[e]	log P[f]	log P_i[f]	log D[g]	log k'[h]	CLOGP[i]	CMR[i]	log(1/MIC)[j]	in vivo activity obsvd[k]	in vivo activity calcd[l]
12	azetidin-1-yl	8.49	9.54		3.264	1.124	2.98	1.62	0.56	0.27	2.385	16.786	3.48	46 (31−72)	15−50
19	$HOCH_2CH_2NH$	8.59	9.62		3.437	1.217	2.47	0.00	−0.97	0.10	1.111	16.653	3.79	10 (6.6−15)	15−50
20	$HOCH_2CH_2NMe$	8.56	9.18		2.970	1.190	2.62	0.00	−0.35	0.16	1.445	17.116	3.50	45	15−50
22	$HOCH_2CMe_2CH_2NH$	8.53	9.72		3.481	1.161	3.03	1.60	0.49	0.43	1.453	18.044	4.12	26 (21−32)	15−50
23	$HOCH_2CH(OH)CH_2NH$[m]	8.55	9.02		2.800	1.180	1.46	0.00	−1.34	0.01	0.286	17.269	3.22	10 (7.3−13)	5−15
25	(2,6-dihydroxy-cyclohexyl)amino[n]	8.56	9.31		3.099	1.189	2.56	0.00	−0.54	0.23	1.411	18.483	3.53	11 (3.5−36)	15−50
27	$Me_2N(CH_2)_3NH$	7.97	8.88	10.13	4.887	2.157	1.91	2.10	−0.06	0.19	1.227	18.260	3.22	5 (3.6−6.2)	15−50
31	Gly-$NMeCH_2CH_2NMe$	6.31	8.10	8.95	2.359	0.809	2.22	0.00	−0.14	−0.10	1.436	19.128	3.54	4 (3−7)	5−15
32	L-Ala-$NMeCH_2CH_2NMe$	6.81	8.13	8.90	2.391	0.891	2.19	0.00	−0.20	0.02	1.745	19.692	4.15	12 (5.8−24)	5−15
34	$Me_2N(CH_2)_3N$(Gly)	7.85	8.74	9.81	4.337	1.927	3.34	1.83	−0.05	−0.25	0.569	19.592	2.96	4 (3.0−6.5)	5−15
35	$Me_2N(CH_2)_3N$(L-Ala)	7.81	8.78	9.60	4.138	1.938	3.55	1.69	−0.08	−0.12	0.878	20.055	3.27	3 (1.9−3.9)	5−15
37	L-Ala-$NMe(CH_2)_3NMe$	7.03	8.53	9.16	3.067	1.307	2.53	0.00	−0.54	−0.16	0.878	20.055	3.27	5 (3.6−6.2)	5−15
39	$Me_2N(CH_2)_2NMe$[o]	5.61	8.19	8.95	2.413	0.863	1.67	0.00	−0.74	−0.23	0.199	22.711	2.44	>50	15−50
41	$Me_2N(CH_2)_3NMe$[o]	7.45	8.43	9.77	3.746	1.376	2.37	0.55	−0.72	−0.80	−0.486	23.175	2.45	41	15−50

[a] Numbering corresponds to that of ref 1. [b] Lowest measured pKa value if dibasic or tribasic, or the sole pKa value if monobasic. [c] Highest measured pKa value if disbasic or middle value if tribasic. [d] Highest pKa in tribasic compounds. [e] See text for details. [f] Calculated from the shift in pKa value in the presence of rapidly stirred 1-octanol. [g] Calculated from measured pKa, log P, and log Pi value(s). [h] Log of capacity factor (see text for details). [i] Calculated from the MEDCHEM program (see ref 7). [j] MIC was determined as described in the Experimental Section of ref 1. [k] For a description of activity class, see text. [l] Calculated from eq 12. [m] Mixture of epimers. [n] Hydroxy groups cis to each other and trans to the amino group. [o] Desmycosin derivatives.

Table 6. Literature pK_a Values Found for the 15 Macrolides Considered in the Article

macrolide	pK_{a1}, pK_{a2}	solvent	method	temp, °C	ref
erythromycin	8.8	not given	not given	not given	19
	8.6−8.9	aqueous	titration	not given	20
	9.1	H_2O:D_2O (9:1)	NMR	30	20
	8.8	D_2O	NMR	not given	21
	8.6	DMF:H_2O (2:1)	not given	not given	22
	8.36	phosphate buffer	pH−solubility profile	37	18
clarithromycin	8.3	DMF:H_2O (2:1)	not given	not given	23
	8.76	phosphate buffer	pH−solubility profile	37	18
erythromycyclamine	8.8, 9.8	DMF:H_2O (2:1)	titration	not given	24
	8.4[a]	DMF:H_2O (2:1)	not given	not given	22
azithromycin	8.85[a]	aqueous	not given	not given	20
rosaramicin	8.4	DMF:H_2O (2:1)	titration	not given	25
oleandomycin	8.5	50% aqueous ethanol	not given	not given	26

[a] The second pK_a was not reported or was undetected in these experiments.

other. The present work attempts to rectify this situation. Among the compounds considered here, log P values of 2.54[12] and 2.91[13] have been noted for erythromycin, and 3.24[13] for clarithromycin. These are similar to the values given in Table 2.

Several groups have reported what are essentially log D values of erythromycin: 1.26,[14] 0.66,[15] 1.62,[16] and 1.70.[17] With the exception of the 0.66 value, these log D's are in good agreement with that given in Table 2. Nakagawa et al.[18] determined partitioning for erythromycin and clarithromycin at 37 °C in the pH range 4.0−8.0. To the best of our knowledge, log P and/or log D values for the other 13 compounds in Table 2 have not yet been published.

Information on pK_a's is not much better. As far as we were able to determine, Table 6 summarizes the literature data available on the pK_a's of the compounds discussed here.[19−26] *The Merck Index*[19] reports a pK_a of 8.8 for erythromycin, but there is no indication of the method, solvent system, temperature, or ionic strength employed. These parameters are critical because they influence the resulting values. For the same compound, variation in pK_a value by using solvents with different dielectric constants can be as much as 0.6 unit.[27−29] Thus, the pK_a values in Table 6 agree in some instances with values reported here, but in other cases are quite different. Because of the different methods and different solvents employed, the literature data cannot be used reliably in QSAR studies owing to these inconsistencies. In the present work, we have used a single method and a single solvent system, so our data are self-consistent and suitable for our purpose.

The current findings have rewarded us with some new insights into those factors governing both the in vitro and in vivo potencies of macrolides. The interpretation of eq 11, which models in vitro potency, is straightforward. For a particular macrolide, the contribution of each term in eq 11 toward potency is shown in Table 7. log P clearly dominates; hence, the negative aspect of log D appears to serve only as a correction to the general lipophilic part of the transport process. The negative sign on the coefficient of the CMR term in eq 11 indicates that the larger macrolides pay a penalty for their size in regard to in vitro potency. Thus, the

Table 7. Contributions of the Various Terms in Eq 11 and the Actual and Predicted log(1/MIC)'s with Residuals

compound	0.724 log *P*	−0.362 log *D*	−0.145 CMR	log(1/MIC) observed	log(1/MIC) estimated[a]	residual
azithromycin	2.909	−0.221	−2.863	4.18	4.001	0.179
carbonate [74211-76-0]	2.113	−0.294	−2.830	3.88	3.165	0.715
erythromycylamine	2.171	0.116	−2.795	3.57	3.667	−0.097
rosaramicin	1.585	−0.308	−2.233	3.46	3.220	0.240
clarithromycin	2.287	−0.569	−2.814	3.28	3.079	0.201
repromicin	1.802	−0.391	−2.211	3.16	3.375	−0.215
erythromycin	2.214	−0.573	−2.747	2.97	3.070	−0.100
OMT	0.716	−0.091	−2.256	2.88	2.546	0.334
tilmicosin	2.750	−0.286	−3.351	2.75	3.288	−0.538
carbomycin A	2.200	−0.949	−3.069	2.73	2.357	0.373
desmycosin	0.724	−0.101	−2.857	2.39	1.941	0.449
carbomycin B	2.547	−1.138	−3.048	2.12	2.537	−0.417
oleandomycin	1.223	−0.156	−2.565	2.04	2.678	−0.638
desmycarosylcarbomycin A	0.217	−0.011	−2.215	1.69	2.167	−0.477
tylosin	1.179	−0.417	−3.370	1.56	1.569	−0.009

[a] The sum of log *P*, log *D*, and CMR columns plus 4.18, the constant term in eq 11.

Table 8. Contributions of the Various Terms in Eq 12 and the Actual and Predicted Activity Classes with Residuals

compound	$-2.321pK_{a(high)}$	1.255p*NF*	−1.241 log *k′*	0.824 log(1/MIC)	act. class actual	act. class estimated[a]	residual
azithromycin	−21.930	4.280	−0.472	3.445	2.5	2.615	−0.115
carbonate [74211-76-0]	−21.373	3.478	−0.422	3.198	2.0	2.172	−0.172
erythromycylamine	−23.091	5.174	0.583	2.943	3.0	2.900	0.100
rosaramicin	−20.399	1.766	−0.782	2.852	1.0	0.729	0.271
clarithromycin	−20.863	2.009	−0.782	2.704	1.0	0.360	0.640
repromicin	−20.492	1.815	−1.067	2.605	0.0	0.152	−0.152
erythromycin	−20.608	1.875	−0.434	2.448	0.0	0.573	−0.573
OMT	−19.494	1.307	0.211	2.374	2.0	1.689	0.311
tilmicosin	−22.186	3.774	−0.323	2.267	1.0	0.824	0.176
carbomycin A	−17.660	0.526	−1.837	2.250	0.0	0.570	−0.570
desmycosin	−19.401	1.261	−0.211	1.970	1.0	0.911	0.089
carbomycin B	−17.521	0.479	−2.185	1.747	0.0	−0.187	0.187
oleandomycin	−20.515	1.826	−0.012	1.682	0.0	0.272	−0.272
desmycarosylcarbomycin A	−19.586	1.353	0.050	1.393	0.0	0.501	−0.501
tylosin	−17.939	0.624	−0.844	1.286	1.0	0.418	0.582

[a] The sum of $pK_{a(high)}$, p*NF*, log *k′*, and log(1/MIC) columns plus 17.29, the constant term in eq 12.

smaller, more hydrophobic macrolides should exhibit the best in vitro potencies.

The interpretation of the model for in vivo activity, eq 12, is more challenging. For a particular macrolide, the contribution of each term in eq 12 toward potency is shown in Table 8. Taking the easier terms first, we observe (as expected) that a low MIC makes a positive contribution to in vivo potency. Also, as noted in our previous paper,[1] those compounds with the smaller log *k*'s (the more polar) in a standard reverse phase HPLC system are the more potent in the mouse *P. multocida* infection model. In keeping with the idea that the more ionized the molecule the less likely it will be absorbed by various biological membranes, eq 12 suggests that compounds with the higher pK_a's tend to be less potent, although this is not obvious by inspection of Table 3. This is because high pK_a's also contribute to high p*NF*'s.

The main difficulty is to reconcile increasing potency with rising p*NF* values. As already noted, the larger the p*NF* value, the smaller the fraction of the neutral species. Thus, we have the paradoxical situation where the less we have of something the more effective it is. Obviously, an alternate interpretation is required. One such comes from considering the make-up of the equations defining p*NF* (eqs 5 and 6). p*NF1* depends on a single pK_a, but p*NF2* depends on two, which leads to much higher values. As shown in Table 8, the p*NF* term contributes to the potencies of monobases over a range 0.479−2.009, but those for dibases covers the range 3.478−5.174. Hence, it appears that it is the number of basic sites that is important in determining in vivo potency. That is, the p*NF* term is important primarily because it is an indirect reflection of the number of basic sites present. Secondarily, differences within these two groups may represent corrections to the negative impact of the pK_a term.

While the chemical structures of macrolides are undoubtedly crucial to their intrinsic antibacterial activities, variations in their potencies appear largely to be functions of their physical properties. To obtain a macrolide with high in vivo potency, there have to be some trade-offs. We need highly lipophilic compounds for low MIC's, but require more polar compounds to achieve good pharmacokinetics. High pK_a's result in large p*NF*'s which contribute positively to in vivo potency, but they also contribute to lower potency as indicated by eq 12. These counterbalancing effects are reflected in Table 8. Erythromycylamine is the most potent among these macrolides. This is owing to its having the highest p*NF*, the lowest log *k′*, and a low MIC. Tylosin is not potent in vivo; with an MIC of 25 μg/mL it is surprising that it is active at all. Its in vivo activity is mainly due to its having one of the lower pK_a's. Perhaps a more telling comparison is that between erythromycylamine 11,12-carbonate and OMT, PD_{50}'s of 30 and 28 mg/kg, respectively. On the plus side, the carbonate's potency gains by having a low MIC (0.10 μg/mL) and a relative large p*NF* (2.8), but it loses by its only moderate log *k′* (0.34) and its high $pK_{a(high)}$ (9.2). In contrast, OMT has a poor MIC (0.78 μg/mL) and a poor p*NF* (1.0), but these are compensated by its having a low log *k′* (−0.17) and a low $pK_{a(high)}$ (8.4). As

one examines Table 8, one sees that nearly every compound's potency appears to be determined by a unique weighting of the factors considered.

This exercise has demonstrated that the in vivo potencies of macrolides can be understood largely in terms of their physical properties. It shows that there are a variety of ways that in vivo activity can be obtained. Certain combinations of properties should give macrolides even greater in vivo potency. For example, to increase p*NF* (a positive influence on potency), it is only necessary to add a third basic site. This is demonstrated by examples in Table 5. Indeed, several, e.g. **35**, meet the criterion for activity class 4 in Table 3. Alternatively, one could use β-hydroxyalkylamines to furnish more polar products with basic sites having lower pK_a's. This in turn should lead to greater potency. Again this is realized by several examples in Table 5, namely, compounds **19**, **23**, and **25**.

Experimental Section

Potentiometric Determination of pK_a and log *P*. Determinations of pK_a's and log *P*'s were performed by the Sirius PCA 101 Potentiometric System. All macrolides were soluble in water at pH 6 or lower. An approximately 1 mM solution of each macrolide at a constant ionic strength of 0.167 N NaCl was titrated from a low to a high pH. The acid used was 0.5 N HCl, and the base, 0.5 N NaOH. The acid and base were standardized to four decimal places using NIST traceable standards. HPLC grade 1-octanol was obtained from the Aldrich Chemical Co. and was water-saturated. The pK_a's and log *P*'s were determined in triplicate under an argon atmosphere at a constant temperature of 25 °C. The values reported here have an average standard deviation of ±0.07 for the pK_a and ±0.09 for the log *P*.

High-Performance Liquid Chromatography (HPLC). Retention times were determined on a YMC 5 μm C-8 column (4.6 mm i.d. × 250 mm length) from Eicon Scientific (P.O. Box 70, Medway, MA). The eluent was a 35:65 (v:v) mixture of MeCN−aqueous (50 mmol) NH_4OAc. The column was maintained at room temperature, and the flow rate at 1.0 mL/min. Each sample was dissolved in the premixed eluent solvent (1 mg/mL), and was injected via an LDC 713 Autosampler (Thermo Separation Products, 3661 Interstate Park Road North, Riviera Beach, FL) into an air actuated Rheodyne 7126 (Cotati, CA) injection valve equipped with a (20 μL) injection loop. The pumping system was an LDC CM 4000. Peaks corresponding to the sample input were detected by UV spectroscopy at 254 nm (carbomycin A, desmycarosylcarbomycin A, rosaramicin) or at 280 nm (carbomycin B, desmycosin, OMT, repromicin, tilmicosin, tylosin) with an LDC SM 3100 UV detector, or with an LDC Refractomonitor IV (azithromycin, clarithromycin, erythromycin, erythromycylamine, erythromycylamine 11,12-carbonate, and oleandomycin).

Acknowledgment. Heather L. Burris, Andrea Davis, Daniel L. Earley, David A. Koss, Mary A. LeMay, Stephen F. Petras, Scott B. Seibel, Annette M. Silvia, and Matthew Teague made valuable technical contributions to this work. We thank them for their fine effort.

References

(1) McFarland, J. W.; Hecker, S. J.; Jaynes B. H.; Jefson, M. R.; Lundy, K. M.; Vu, C. B.; Glazer, E. A.; Froshauer, S. A.; Hayashi, S. F.; Kamicker, B. J.; Reese, C. P.; Olson, J. A. Repromicin Derivatives with Potent Antibacterial Activity against *Pasteurella multocida*. *J. Med. Chem.* **1997**, *40*, 1041−1045.

(2) Avdeef, A. pH-Metric log *P*. Part II: Refinement of Partition Coefficients and Ionization Constants of Multiprotic Substances. *J. Pharm. Sci.* **1993**, *82*, 183−190.

(3) Avdeef, A. pH-Metric log P. Part I. Difference Plots for Determining Ion-Pair Octanol-Water Partition Coefficients of Multiprotic Substances. *Quant. Struct.-Act. Relat.* **1992**, *11*, 510−517.

(4) Slater, B.; McCormack, A.; Avdeef, A.; Comer, J. E. A. pH-Metric log *P*. 4. Comparison of Partition Coefficients Determined by HPLC and Potentiometric Methods to Literature Values. *J. Pharm. Sci.* **1994**, *83*, 1280−1283.

(5) Taylor, P. J. Hydrophobic Properties of Drugs. In *Comprehensive Medicinal Chemistry*; Ramsden, C. A., Ed.; Pergamon Press: Oxford, 1990; Vol. 4, pp 241−294.

(6) Horváth, C.; Melander, W.; Molnár, I. Liquid Chromatography of Ionogenic Substances with Nonpolar Stationary Phases. *Anal. Chem.* **1977**, *49*, 142−154.

(7) MEDCHEM, version 3.54, Daylight Chemical Information Systems, Inc., 18500 Von Karman Ave, Suite 450, Irvine, CA 92715.

(8) Release 4.4.1 of RS/1, Bolt, Beranek and Newman, 150 Cambridge Park Dr, Cambridge, MA 02140.

(9) Modde for Windows, User's Guide to Modde, version 2.1, Umetri AB, Box 1456, S-901 24 Umeå, Sweden, 1994.

(10) Dearden, J. C. Molecular Structure and Drug Transport. In *Comprehensive Medicinal Chemistry*; Ramsden, C. A., Ed.; Pergamon Press: Oxford, 1990; Vol. 4, pp 375−411.

(11) Kirst, H. A. Semi-synthetic Derivatives of Erythromycin. In *Progress in Medicinal Chemistry*; Ellis, G. P., Luscombe, D. K., Eds.; Elsevier Science Publishers B. V.: Amsterdam, 1993; Vol. 30, pp 57−88.

(12) Martin, Y. C.; Jones, P. H.; Perun, T. J.; Grundy, W. E.; Bell, S.; Bower, R. R.; Shipkowitz, N. L. Chemical Modification of Erythromycin Antibiotics. 4. Structure-Activity Relationships of Erythromycin Esters. *J. Med. Chem.* **1972**, *15*, 635−638.

(13) Goldman, R. C.; Zakula, D.; Flamm, R.; Beyer, J.; Capobianco, J. Tight Binding of Clarithromycin, Its 14-(*R*)-Hydroxy Metabolite, and Erythromycin to *Helicobacter pylori* Ribosomes. *Antimicrob. Agents Chemother.* **1994**, *38*, 1496−1500.

(14) Selassie, C. D.; Hansch, C.; Khwaja, T. Structure-Activity Relationships of Antineoplastic Agents in Multidrug Resistance. *J. Med. Chem.* **1990**, *33*, 1914−1919.

(15) Wildfeuer, A.; Lemme, J.-D. Zur Pharmakokinetik von Josamycin. (On the Pharmacokinetics of Josamycin.) *Arzneim.-Forsch.* **1985**, *35*, 639−643.

(16) Ohno, M.; Ohta, K.; Morishita, M. Physicochemical properties of rokitamycin. *Iyakuhin Kenkyu* **1987**, *18*, 634−639 [cited by Muto, Y.; et al. *Antimicrob. Agents Chemother.* **1989**, *33*, 242−244, and by Kobayashi, Y.; et al. *FEMS Microbiol. Lett.* **1991**, *81*, 141−144].

(17) Manuel, C.; Dellamonica, P.; Rosset, M. J.; Safran, C.; Pirot, D.; Audegond, L.; Pechere, J. C. Penetration of roxithromycin (ROX) into brain tissue (BT). *Abstracts of the 28th Intersci. Conf. Antimicrob. Agents Chemother.* **1988**, 328 (No. 1224).

(18) Nakagawa, Y.; Itai, S.; Yoshida, T.; Nagai, T. Physiochemical Properties and Stability in the Acidic Solution of a New Macrolide Antibiotic, Clarithromycin, in Comparison with Erythromycin. *Chem. Pharm. Bull.* **1992**, *40*, 725−728.

(19) Erythromycin. In *The Merck Index, Eleventh Edition*; Budavari, S., Ed.; Merck & Co.: Rahway, NJ, 1989; pp 577−578 (No. 3626).

(20) Goldman, R. C.; Fesik, S. W.; Doran, C. C. Role of Protonated and Neutral Forms of Macrolides in Binding to Ribosomes from Gram-Positive and Gram-Negative Bacteria. *Antimicrob. Agents Chemother.* **1990**, *34*, 426−431.

(21) Gharbi-Benarous, J.; Delaforge, M.; Jankowski, C. K.; Girault, J.-P. A Comparative NMR Study between the Macrolide Antibiotic Roxithromycin and Erythromycin A with Different Biological Properties. *J. Med. Chem.* **1991**, *34*, 1117−1125.

(22) Kobrehel, G.; Tamburasev, Z.; Djokic, S. Erythromycin series. IV. Thin-layer chromatography of erythromycin, erythromycin oxime, erythromycylamine and their acyl derivatives. *J. Chromatogr.* **1977**, *133*, 415−419.

(23) Delaforge, M.; Ladam, P.; Bouille, G.; Benarous, J. G.; Jaouen, M.; Girault, J. P. pH Effects on the N-Demethylation and Formation of the Cytochrome *P*-450 Iron II Nitrosoalkane Complex for Erythromycin Derivatives. *Chem.-Biol. Interact.* **1992**, *85*, 215−227.

(24) Massey, E. H.; Kitchell, B. S.; Martin, L. D.; Gerzon, K. Antibacterial Activity of 9(S)-Erythromycylamine−Aldehyde Condensation Products. *J. Med. Chem.* **1974**, *17*, 105−107.

(25) Kishi, T.; Harada, S.; Yamana, H.; Miyake, A. Studies on Juvenimicin, a New Antibiotic. II. Isolation, Chemical Characterization and Structures. *J. Antibiot.* **1976**, *29*, 1171−1181.

(26) Els, H.; Celmer, W. D.; Murai, K. Oleandomycin (PA-105). II. Chemical Characterization (I). *J. Am. Chem. Soc.* **1958**, *80*, 3777−3782.

(27) Shedlovsky, T.; Kay, R. L. The ionization constant of acetic acid in water-methanol mixtures at 25 °C from conductance measurements. *J. Phys. Chem.* **1956**, *60*, 151−155.

(28) Shedlovsky, T. The behavior of carboxylic acids in mixed solvents. In *Electolytes*, Pesce, B., Ed.; Pergamon Press: New York, 1962; pp 146−151.

(29) Yasuda, M. Dissociation constants of some carboxylic acids in mixed solvents. *Bull. Chem. Soc. Jpn.* **1959**, *32*, 429−432.

JM960436I